CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION, CFSTI
DOCUMENT MANAGEMENT BRANCH 410.11

LIMITATIONS IN REPRODUCTION QUALITY

Accession # AD 606 371

- [x] 1. We regret that legibility of this document is in part unsatisfactory. Reproduction has been made from best available copy.
- [ ] 2. A portion of the original document contains fine detail which may make reading of photocopy difficult.
- [ ] 3. The original document contains color, but distribution copies are available in black-and-white reproduction only.
- [ ] 4. The initial distribution copies contain color which will be shown in black-and-white when it is necessary to reprint.
- [ ] 5. Limited supply on hand; when exhausted, document will be available in Microfiche only.
- [ ] 6. Limited supply on hand; when exhausted document will not be available.
- [ ] 7. Document is available in Microfiche only.
- [ ] 8. Document available on loan from CFSTI (TT documents only).
- [ ] 9.

Processor: jat

NBS____ 9/64

ON THE COMPUTATIONAL SOLUTION OF DYNAMIC—PROGRAMMING PROCESSES—I ON A TACTICAL AIR—WARFARE MODEL OF MENGEL

By

Richard Bellman
Stuart Dreyfus

P—1072

Revised

May 23, 1957

The RAND Corporation
1700 MAIN ST • SANTA MONICA • CALIFORNIA

SUMMARY

This is the first of a series of papers devoted to the computational solution of dynamic programming processes. In it we use the functional-equation approach to treat a tactical air-warfare model that A. Mengel previously has considered by means of classical variational techniques.

CONTENTS

ON THE COMPUTATIONAL SOLUTION OF DYNAMIC-PROGRAMMING PROCESSES-I ON A TACTICAL AIR-WARFARE MODEL OF MENGEL

1. INTRODUCTION

This is the first of a series of papers devoted to the computational solution of dynamic-programming processes. Although the papers are linked together by a common method each of the diverse problems we shall treat possesses particular features of interest and difficulty that make a detailed exposition of the coding worthwhile.

It is planned eventually to present all the papers of the series in the form of a book.

We would like to express our appreciation to E. W. Paxson for a number of helpful comments and suggestions which we have incorporated in the paper.

2. ATTRITION PROCESSES

The study of attrition processes arising from military campaigns leads to a class of variational problems that are particularly well suited to dynamic programming.

Consider the following model. Let the state of Blue's forces at time t be specified by the vector x, with components $x_1, x_2, \ldots, x_M$, and the state of Red's forces be specified by y, with components $y_1, y_2, \ldots, y_N$. At each stage of the process, which may be discrete or continuous—and this has less to do with reality than with the type of computing machine which is available, a digital computer or an analog computer—each side allocates a certain portion of the forces to combat, obtaining

in this way a certain payoff and suffering, in return, a certain attrition. Let z be the allocation vector of Blue and w the allocation vector of Red. The natural constraints are, in vector form,

(1) $$0 \leq z \leq x, \quad 0 \leq w \leq y;$$

that is,

(2) $$0 \leq z_i \leq x_i, \quad i = 1,2,\ldots,M,$$
$$0 \leq w_j \leq y_j, \quad j = 1,2,\ldots,N.$$

The single-stage payoff is determined as some function

(3) $$R(x,y,z,w)$$

(in practice, usually the most difficult function to decide upon) and we assume that we know the attrition due to combat, so that

(4) $$\frac{dx}{dt} = F(x,y,z,w), \; x(0) = q_1$$
$$\frac{dy}{dt} = G(x,y,z,w), \; y(0) = q_2$$

where q_1 and q_2 are the initial forces.

The mathematical problem is then that of determining

(5) $$\min_w \quad \max_z \int_0^T R(x,y,z,w)dt,$$

where T is the duration of the process, subject to the constraints (2) and the relations (4).

Alternatively, we may wish to determine

(6) $$\max_z \quad \min_w \int_0^T R(x,y,z,w)dt.$$

Since, in general, the determination of min max or max min is a formidable problem, particularly if min max $\neq$ max min, we shall reduce the magnitude of the problem by fixing Red's strategy, say $w = w^*$, and proceeding to determine

$$(7) \qquad \max_z \int_0^T R(x,y,z,w^*)dt,$$

subject to

$$(8) \qquad (a) \quad 0 \leq z \leq x,$$

$$(b) \quad \frac{dx}{dt} = F(x,y,z,w^*), \quad x(0) = q_1,$$

$$\frac{dy}{dt} = G(x,y,z,w^*), \quad y(0) = q_2.$$

Problems of this type are difficult, using conventional methods, because of the presence of the constraints, and the analytic structure of the functions F, G and R.

3. MENGEL'S MODEL

Let us now consider the attrition process that has been discussed by Arnold Mengel, using classical variational techniques [5].

Considering only air forces consisting of one type of plane, for the purpose of an exploratory model, we have the equations(1)

$$(1) \qquad \dot{x} = G_1(x,y,s_2) = r_1 - a_2x - x(1 - e^{-b_2 s_2 y/x}),$$

$$\dot{y} = G_2(x,y,s_1) = r_2 - a_1y - y(1 - e^{-b_1 s_1 x/y}),$$

---

(1) r_1, r_2 are replacement rates of new aircraft. The terms a_2x, a_1y represent operational (non-combat) attrition rates. The term $s_2y(=w)$ is Red's counter air effort over $(t, t+dt)$, with "kill potential" b_2s_2y. The chance any one of the x aircraft is killed by one of the s_2y attacks is b_2. The average number of attacks per Blue aircraft is $n = s_2y/x$. Hence the probability of survival is $(1-b_2)\, n \cong \exp(b_2n)$.

for fixed s_2, with the payoff function(1)

(2) $$J(s_1) = \int_0^T \left[(1-s_1)x-(1-s_2)y\right]dt.$$

Here

(3) $x(t)$ = the number of Blue aircraft at time t,

$y(t)$ = the number of Red Aircraft at time t.

The allocation variables are

(4) $s_1(t)$ = fraction of Blue sorties on counter-air strikes,

$s_2(t)$ = fraction of Red sorties on counter-air strikes.

As mentioned above, we shall fix $s_2(t)$, in this case by assuming various constant levels, s_2^*, and then maximizing $J(s_1)$ over all $s_1(t)$ satisfying $0 \leq s_1(t) \leq 1$.

4. DYNAMIC-PROGRAMMING APPROACH—I

Setting

(1) $$\max_{s_1} J(s_1) = f(q_1,q_2,T),$$

we obtain, as in [1], the nonlinear partial differential equation

(2) $$\frac{\partial f}{\partial T} = \max_{0 \leq s_1 \leq 1} \left[(1-s_1)q_1-(1-s_2^*)q_2+G_1(q_1,q_2,s_2)\frac{\partial f}{\partial q_1} + G_2(q_1,q_2,s_1)\frac{\partial f}{\partial q_2}\right]$$

with

(3) $$f(q_1,q_2,0) = 0.$$

This equation may be solved numerically using approximating difference equations in the usual fashion. In practice, we en-

---

(1) This measures the total excess of Blue's combat capability over Red's during the campaign on missions other than counter-air.

countered a great deal of difficulty with this method due to instability arising from transition curves. Consequently, we went over to the method we shall present in the next section. This method has applications to the numerical integration of other types of partial differential equations, a matter which we have discussed elsewhere.

5. DYNAMIC-PROGRAMMING APPROACH—II

Let us consider the following discrete process. Divide the interval $[0,T]$ into N equal parts of length Δ,

$$0 \quad \Delta \quad 2\Delta \quad \cdots \quad k\Delta \quad \cdots \quad N\Delta = T$$

Let us assume that decisions may be made only at times $k\Delta$, $k = 0,1,2,\ldots,N-1$. As far as actual processes are concerned, this may be a more realistic assumption than that of a continuous process.

In place of the equations (6) of Section 2, we have the difference equations, or recurrence relations,

$$(1) \qquad \begin{aligned} x_{k+1} &= x_k + G_1(x_k, y_k, s_2(k))\,\Delta, & x_0 &= q_1, \\ y_{k+1} &= y_k + G_2(x_k, y_k, s_1(k))\,\Delta, & y_0 &= q_2, \end{aligned}$$

where

$$(2) \qquad \begin{aligned} &x_k = x(k\Delta),\ y_k = y(k\Delta), \\ &s_{1k} = s_1(k\Delta). \end{aligned}$$

The sequence $\{s_{1k}\}$ is to be chosen to maximize

(3) $$J(s_1) = \sum_{k=0}^{N-1} \left[(1-s_{1k})x_k-(1-s_{2k})y_k\right],$$

subject to the restriction $0 \leq s_{1k} \leq 1$.

Let

(4) $$\max J(s_1) = f_N(q_1,q_2),$$

with

(5) $$f_1(q_1,q_2) = q_1 - (1-s_{2k})q_2.$$

The basic recurrence relation used to compute the sequence $\{f_N(q_1,q_2)\}$ is

(6) $$f_N(q_1,q_2) = \max_{0\leq s_1\leq 1} \Big[(1-s_1)q_1 - (1-s_2)q_2 + f_{N-1}(q_1 + G_1(q_1,q_2,s_2)\Delta,\ q_2 + G_2(q_1,q_2,s_1)\Delta)\Big]$$

for $N = 1,2,\ldots$.

6. TIME-DEPENDENT PROCESSES

Let us consider a process of the same general type in which the attrition and payoff functions depend upon time. Thus

(1) $$\frac{dx}{dt} = G_1(x,y,s_2,t), \quad x(0) = q_1,$$

$$\frac{dy}{dt} = G_2(x,y,s_1,t), \quad y(0) = q_2,$$

and we wish to choose s_1 so as to maximize

(2) $$J(s_1) = \int_0^T P(x,y,s_1,t)dt.$$

In this case, we keep the terminal point T fixed and describe the state of the process by means of the resources and the starting

point.

The discrete maximization problem is then: Maximize

(3) $$J_R(s_1) = \sum_{k=R}^{N} P(x_k, y_k, s_{1k}, k) ,$$

subject to the constraints

(4) $$x_{k+1} = x_k + G_1(x_k, y_k, s_{2k}, k), \quad x_R = q_1,$$

$$y_{k+1} = y_k + G_2(x_k, y_k, s_{1k}, k), \quad y_R = q_2.$$

Setting

(5) $$\max \; J_R(s_1) = f_R(q_1, q_2),$$

we obtain the recurrence relations

(6) $$f_R(q_1, q_2) = \max_{0 \le s_1 \le 1} \Big[P(q_1, q_2, s_1, R)$$
$$+ f_{R+1}(q_1 + G_1(q_1, q_2, s_2, R), q_2 + G_2(q_1, q_2, s_1, R)) \Big],$$

for $R = 0, 1, 2, \ldots, N-1$, with

(7) $$f_N(q_1, q_2) = \max_{0 \le x_1 \le 1} \quad P(q_1, s_2, s_1, N).$$

7. DISCUSSION OF COMPUTATIONAL PROCEDURES

The recursive nature of the problem makes it particularly suited to digital computation. A fairly sizable problem can be solved in 500 instructions, leaving the bulk of high-speed storage available for tabulation of the functions. In this initial study the luxury of floating-point arithmetic was allowed, due to uncertainty concerning the ranges of the variables. Considerable additional time and space could be saved in

later studies by fixed-point programming.

The program itself can be divided into 3 logical sections. A master routine does the bookkeeping, tallying, and sequencing; a subroutine evaluates $F(q_1,q_2,s_1) + f_R(q_1+G_1,q_2+G_2)$ for given q_1,q_2,s_1,s_2, and a table of f_R; a second subroutine performs the maximization of the above expression over the interval $0 \leq s_1 \leq 1$.

The flow-chart governing the computation is shown in Figure 8.

In more detail, the computatio· proceeds as follows. Under control of the master, the constants Δq_1, Δq_2, s_2, N, and m+1 are input: Δq_1 and Δq_2 determine the density of the grid over which the function f_R is to be evaluated; the parameter s_2 is Red's constant strategy; N is the number of stages for which the process is allowed to continue; and m+1 determines the size of the grid, its dimensions being $m\Delta q_1$ by $m\Delta q_2$.

The quantity $f_1(q_1,q_2)$ is evaluated over the grid as $\max\limits_{0\leq s_1\leq 1}\left[F(q_1,q_2,s_1) + f_o\ (q_1+G_1,q_2+G_2)\right]$. Since the return from a zero-stage war is identically zero, the maximum always occurs when $s_1 = 0$ so that $f_1(q_1,q_2)$ is merely $q_1-(1-s_2)q_2$. This corresponds to the fact that during the last stage of a war Blue's airpower will be directed entirely against Red's ground forces.

The calculation of f_N with f_{N-1} now known is not quite so trivial However, due to the recurrence relation, this calculation actually defines the remainder of the program. Suppose we wish to evaluate $f_N\ (Q_1,Q_2)$, where (Q_1,Q_2) is some $(i\ \Delta\ q_1, j\Delta q_2)$.

For a particular $s_1 = S_1$ we evaluate $(Q_1 + G_1(Q_1,Q_2,s_2),$ $Q_2 + G_2(Q_1,Q_2,S_1))$. This determines a point X in the q_1,q_2 plane. Now X falls within a rectangle of dimensions Δq_1 by Δq_2, where f_{N-1} is known at the four corners, and $f_{N-1}(X)$ is found by linear interpolation. By adding $(1-S_1)Q_1 - (1-S_2)Q_2$, we determine $f_N(Q_1,Q_2)$ for $s_1 = S_1$. We need only to repeat this process where s_1 takes on values between 0 and 1 to determine the maximum. Since for the majority of a process $s_1 = 0$ or 1, it is expedient first to test for an endpoint maximum before searching the interior region. Since f_R for each R, R=1,2,...,N, is evaluated over a grid of $(m+1)^2$ points, it is essential to optimize the search process. Consequently the technique described in [4] was adopted.

Once the function has been evaluated for a fixed R, R+1 replaces R-1, the newly calculated table f_R replaces f_{R+1} in high-speed storage, and the calculation of f_{R+1} begins.

Since s_1, the Blue strategy which maximizes f, is generally of more interest than the resulting payoff, f, a table of s_1 associated with each f_R is stored and punched out prior to the computation of f_{R+1}.

When R reaches N, assumed length of the conflict, the following information has been obtained:

1) The return attainable by Blue in an N-stage war, where Blue enters the conflict with q_1 planes, Red with q_2, Red uses the fixed allocation between air and ground support s_2, and Blue uses an optimal allocation. This by definition is $f_N(q_1,q_2)$.

2) Blue optimal strategy during the first stage of the N-stage war. This is the s_1 which maximizes f_N.

3) Blue optimal strategy during the first stage of a war of duration R, R=1,2,...,N-1, for any initial forces q_1 and q_2. These tables are punched during the stage-by-stage calculation.

The process determining explicitly an optimal policy is essentially that described above, but in reverse. Knowing R, the stage, and (q_1,q_2), we refer to the table of s_1's to determine the strategy associated with (q_1,q_2). Employing this strategy we find ourselves at the $(R+1)^{st}$ stage and Blue possesses q_1+G_1 planes against Red's q_2+G_2. Furthermore, evaluation of $(1-s_1)q_1 + (1-s_2)q_2$ produces the payoff during the $(R)^{th}$ time interval. By referring to the tables of optimal strategies for the $(R+1)^{st}$ stage and initial forces q_1+G_1 and q_2+G_2, we determine the optimal allocation for the $(R+1)^{st}$ period. The process is one of repeatedly determining the initial strategy in wars of decreasing length and decreasing forces. The sequence defines Blue's optimal policy.

At this time, only the most essential computational variations have been investigated. For example, it was found that due to the near-linear behavior of f_R over the range 0 to 10000, $\Delta q_1 = \Delta q_2 = 500$ gave sufficiently accurate results to justify its use. This same property led to the choice of linear interpolation throughout the grid. Two versions of the Fibonaccian search method of S. Johnson [4] were considered, one using predetermined points of evaluation, the other calculating the points.

The latter, of course, is more general, but the advantages offered by less calculation and faster convergence led to the choice of the former. Printout of all functions and strategy values along the grid was made optional. Considerable time was saved by suppression of printing when s_1 equaled 0, a frequent occurrence during the initial phases of a calculation.

8. GRAPHS

Results are shown on the following pages. Figure 1 shows the changing relative strengths of the rival air forces when Blue employs an optimal policy. Due to Blue's initial counter-air tactics, Red's force is reduced during the early stages, while Blue's force drops suddenly in the later stages when counter-ground strategy is used. Figure 2 shows that Blue's initial allocation of planes is against the Red air force if the total number of planes in each force is fairly even. If Blue has a marked numerical superiority or inferiority, a counter-ground strategy should be employed. Figure 3 depicts Blue's strategy where initial forces are equal. The next three graphs show Blue's strategy as it changes with time for all initial conditions. Figure 7 shows the excess sorties flown by Blue as a result of employing an optimal, rather than constant, strategy. The use of an optimal policy in this particular example is shown to be equivalent to about 800 planes; i.e., with the given parameter values Blue can start with 800 fewer planes and still fly as many sorties as Red during a 15-stage conflict.

Throughout these numerical examples, we have fixed both sides replacement rates, r_1 and r_2, as 100 planes per stage, non-combat

attrition rates, a, as .1, and kill probabilities, b, as .2. Figure 8 shows the flow diagram used for computation.

This structure of the maximizing strategy is typical of a large class of problems. The reason for it lies in the concavity of the function appearing in (5.6) as a function of s_1. This concavity in turn is based upon the linearity of the pay-off function, and the concavity of the attrition function. Two conclusions can be drawn from this. In the first place, it shows that great mathematical simplifications ensue when we introduce concave functions, or convex functions if we are minimizing. Thus, if we have functions which for one reason or another are not concave, it may be well initially to use concave approximations to these functions. On the other hand, these results show the dangers inherent in mathematical models. In the real world, such concentration on counter-air or counter-surface at various phases of the campaign is dubious. Catastrophic loss by ground forces might occur before the time T of the campaign has elapsed. Such effects are not measured by the "uniform" pay-off function J.

There are, of course, dangers in conclusions based upon this one-sided approach in which we fix Red's strategy. Iteration procedures may be considered in which we alternately fix one side's policy and then the other's. These must be used with care, since we know from much simpler games that unless some feedback from stage to stage is used, the results will not converge.

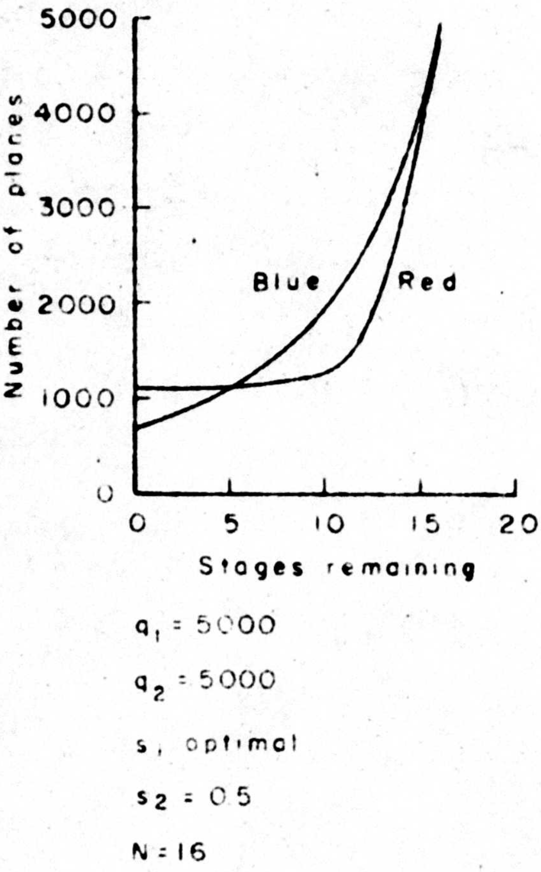
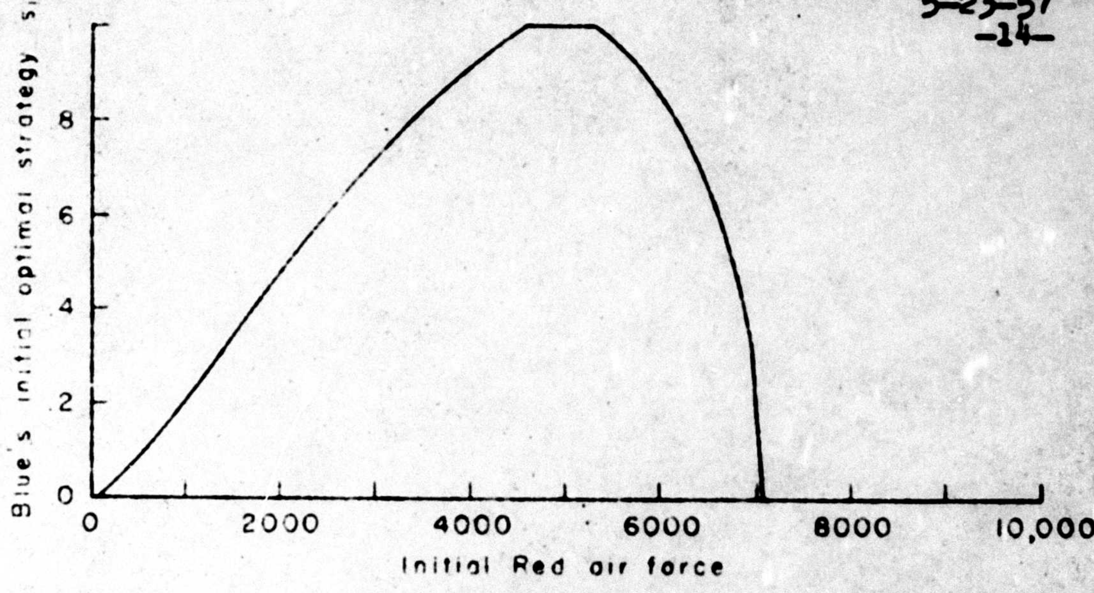
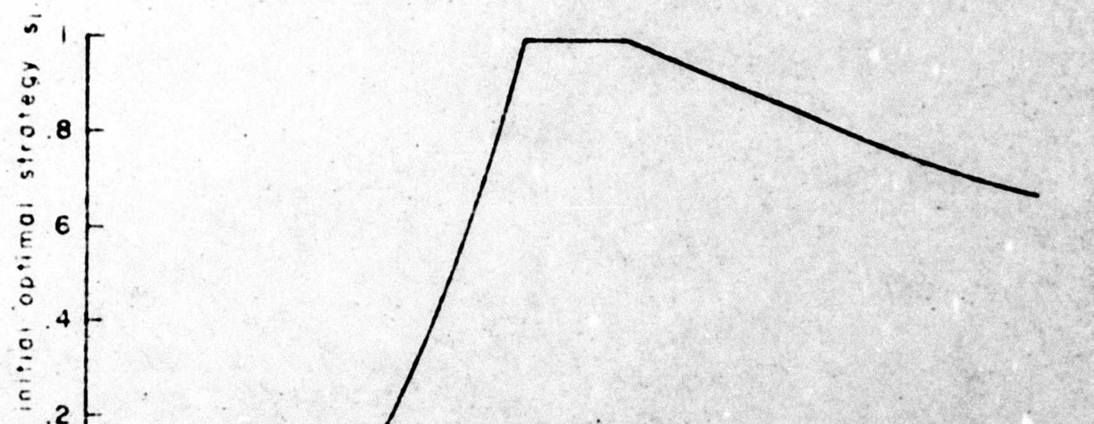


Fig. 1—Changing relative strengths of Blue and Red air forces under optimal Blue policy.

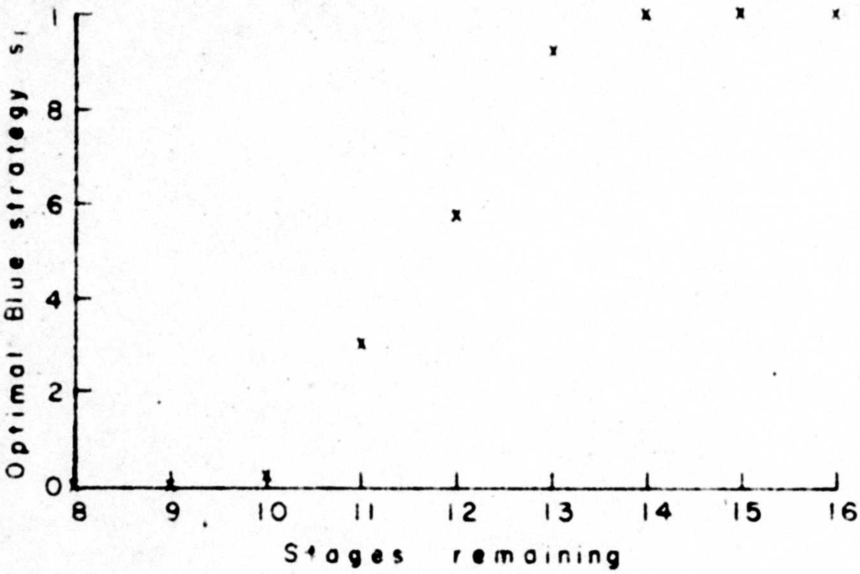


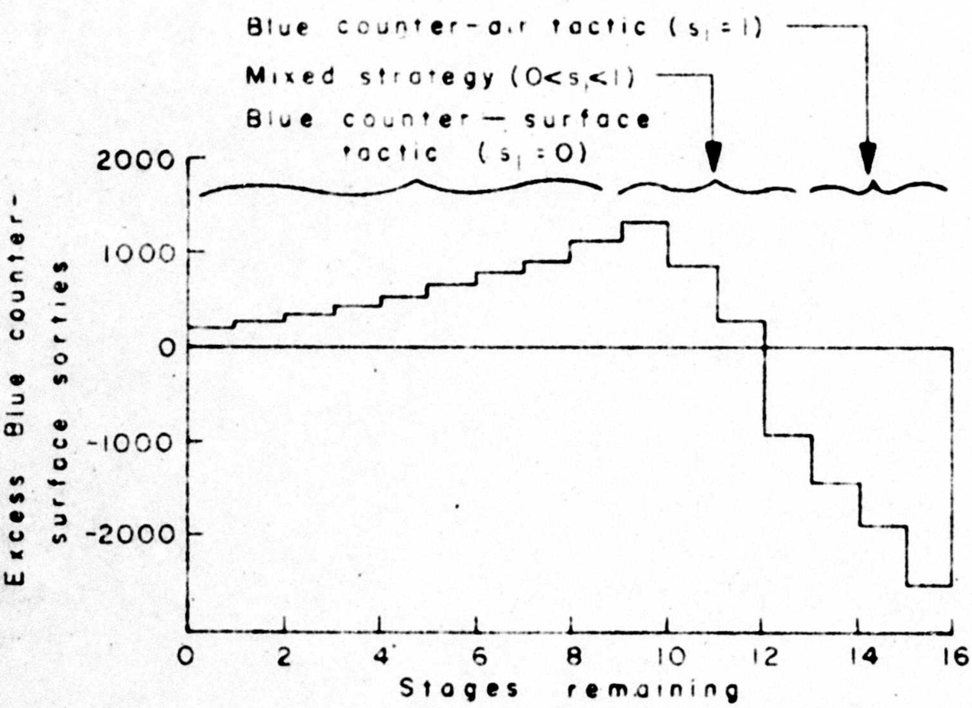
Initial Blue air force = 5000
12-stage war
s_2 = .5



Initial Red air force = 5000
12 stage war
s_2 = 5

Fig. 2—Initial allocation of Blue planes under varying relative air forces.





$s_2 = 5$

$q_1 = 5000$ initially

$q_2 = 5000$ initially

Fig. 3—Blue's strategy when initial forces are equal.

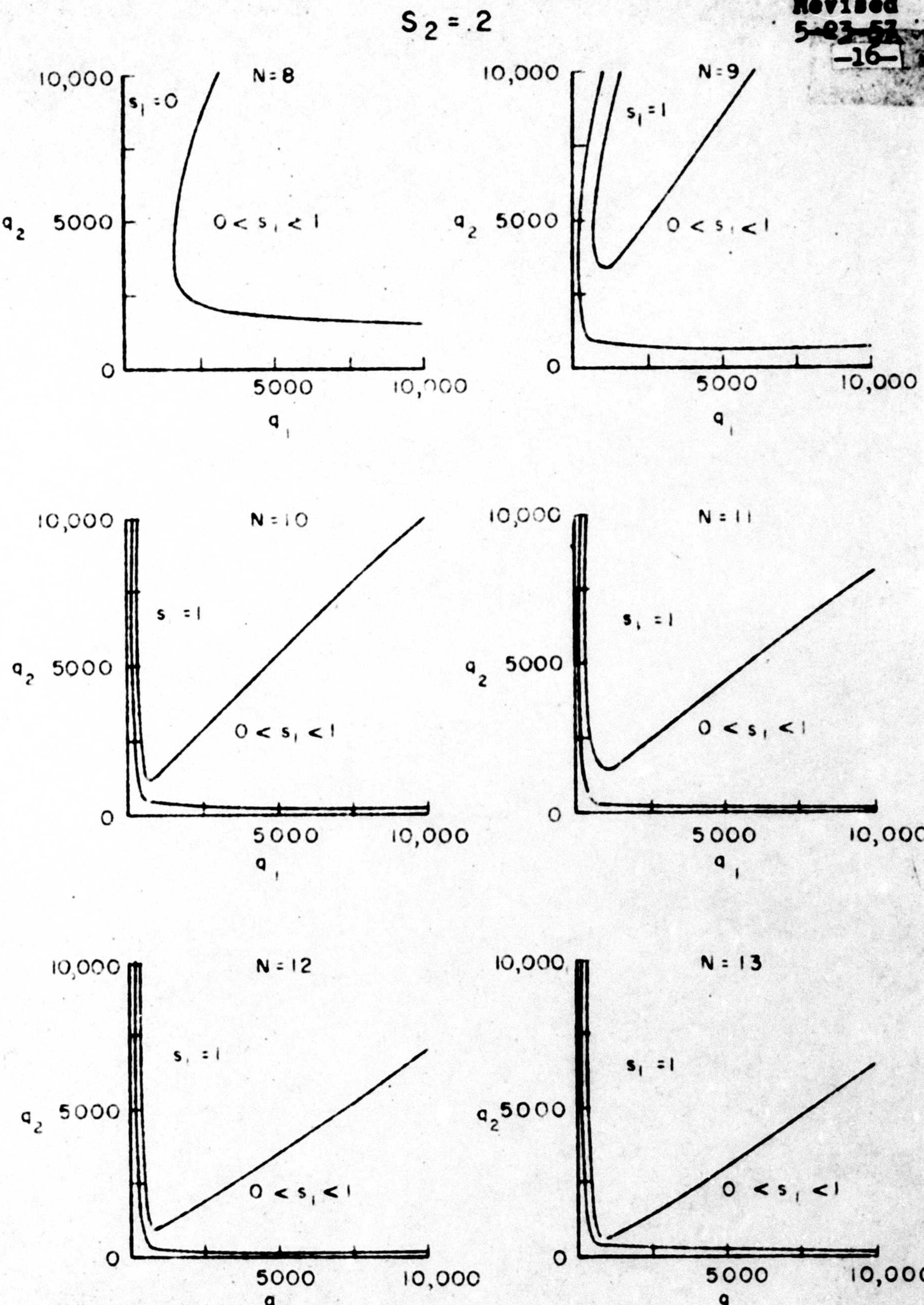


Fig. 4—Variation of Blue's strategy s_1 with time (N=8 to 13) for initial Blue force q_1 and initial Red force q_2 between 0 and 10,000 and for Red strategy $s_2 = 0.2$.

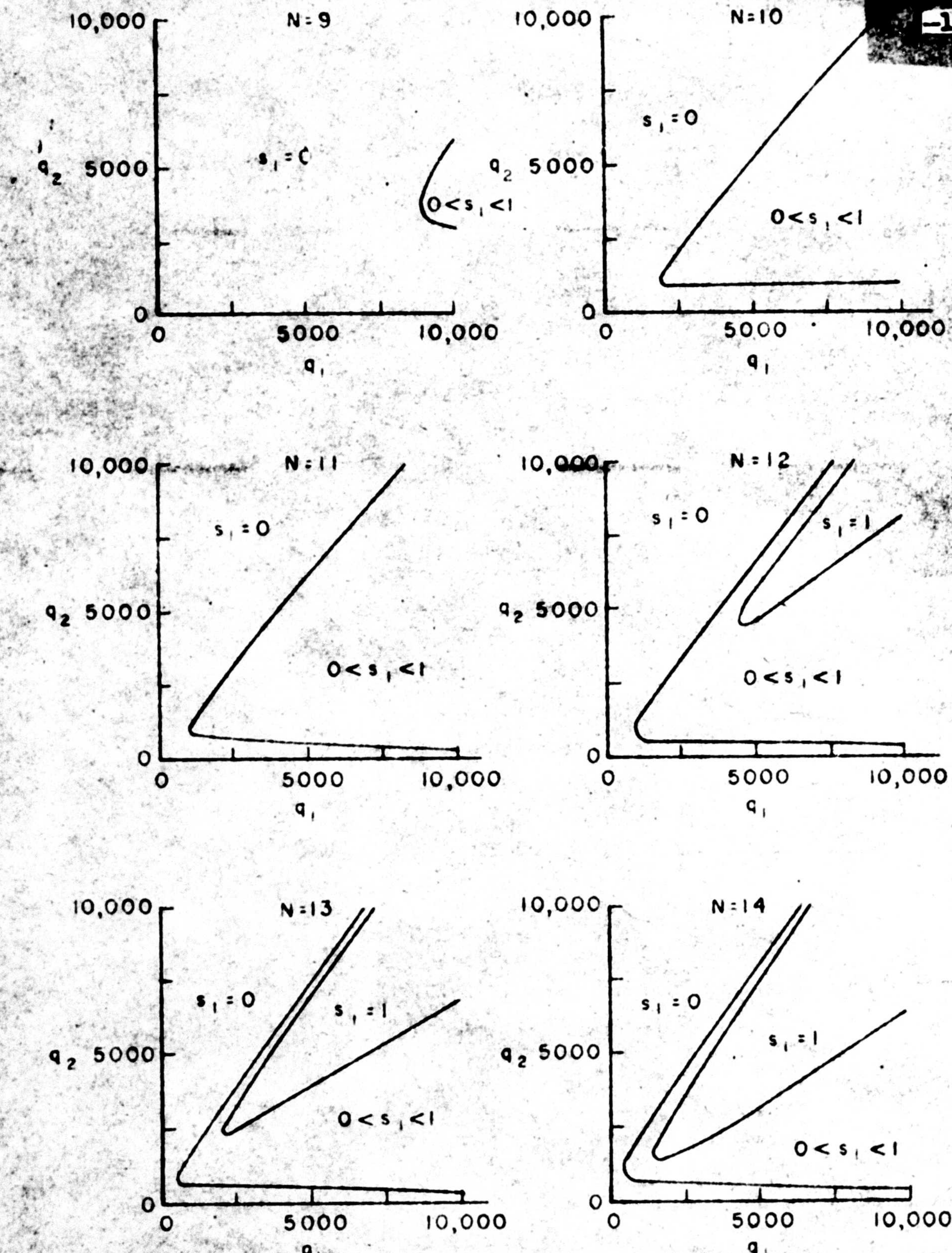


Fig. 5—Variation of Blue's strategy s_1 with time (N=9 to 14) for initial Blue force q_1 and initial Red force q_2 between 0 and 10,000 and for Red strategy $s_2 = 0.5$.

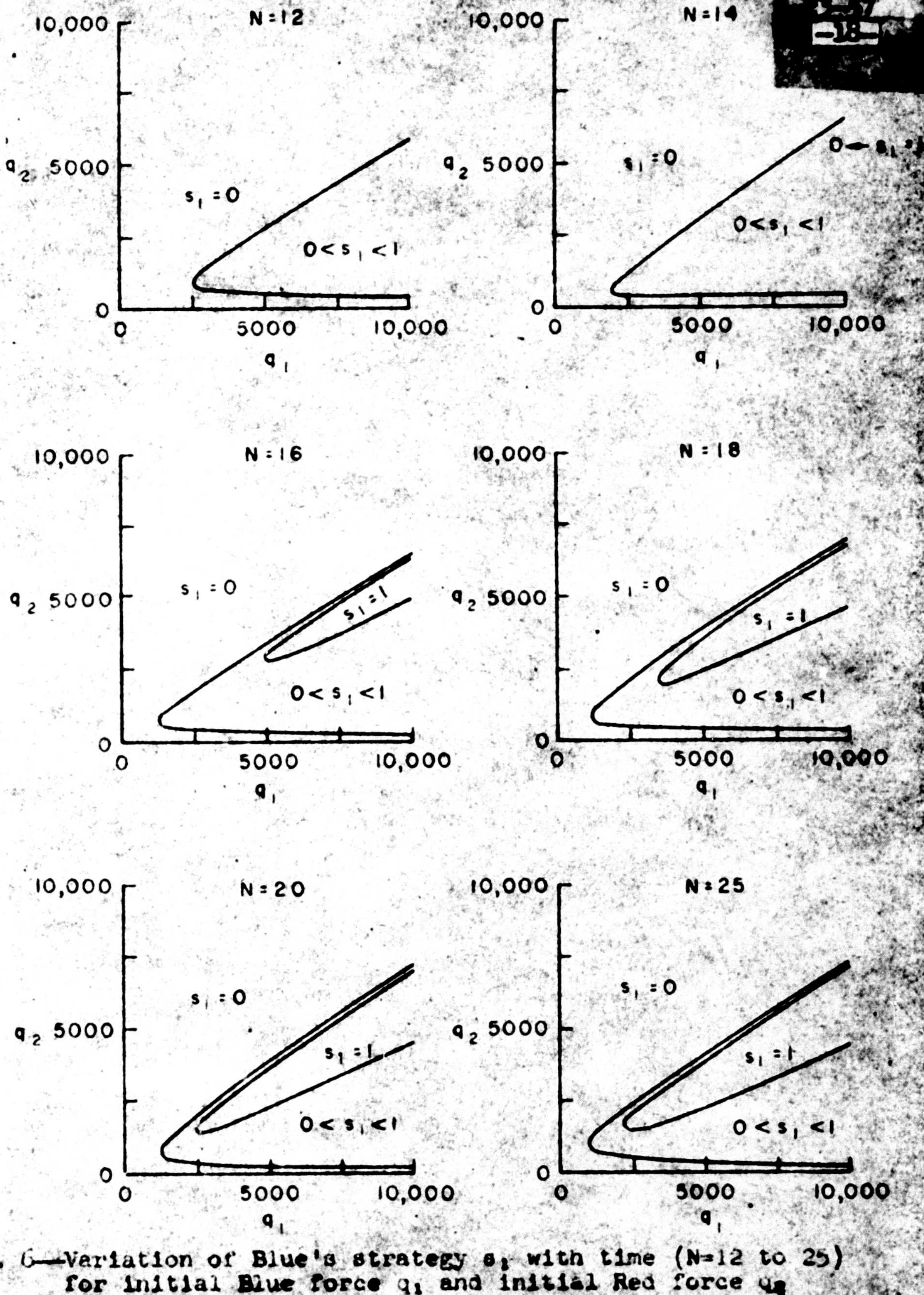
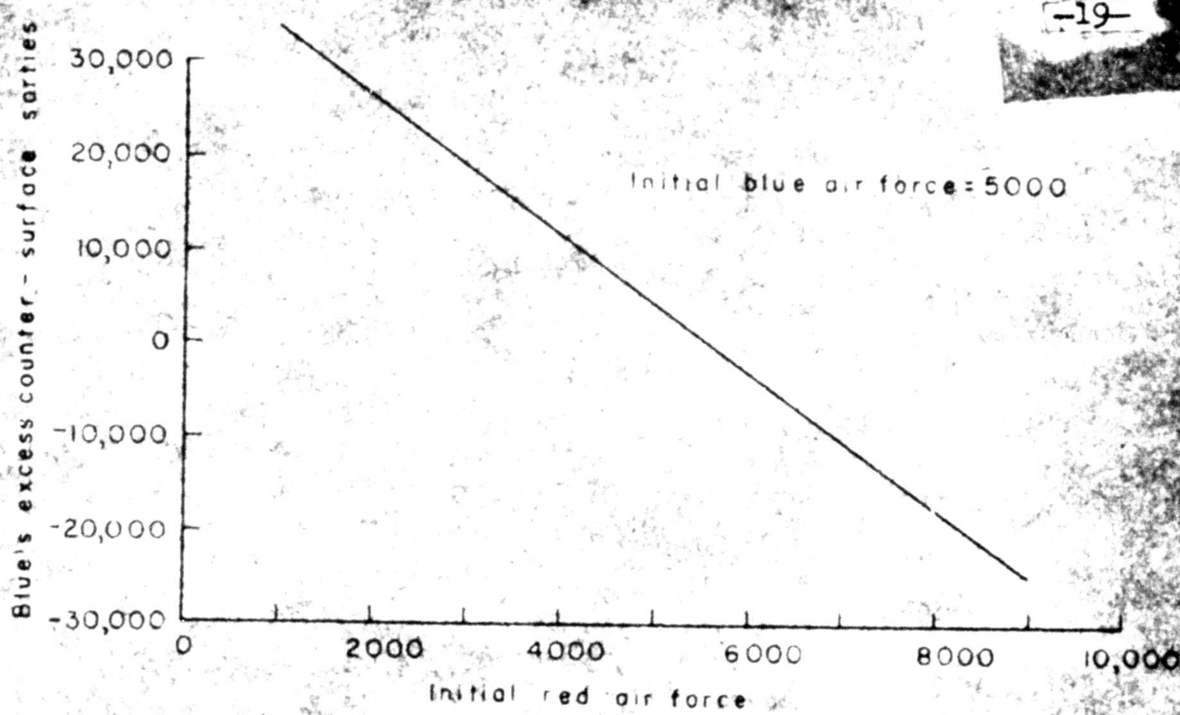


Fig. 6—Variation of Blue's strategy s_1 with time (N=12 to 25) for initial Blue force q_1 and initial Red force q_2 between 0 and 10,000 and for Red strategy s_2 = 0.8.



15 stage war
Blue uses optimal strategy
Red strategy $S_2 = .5$

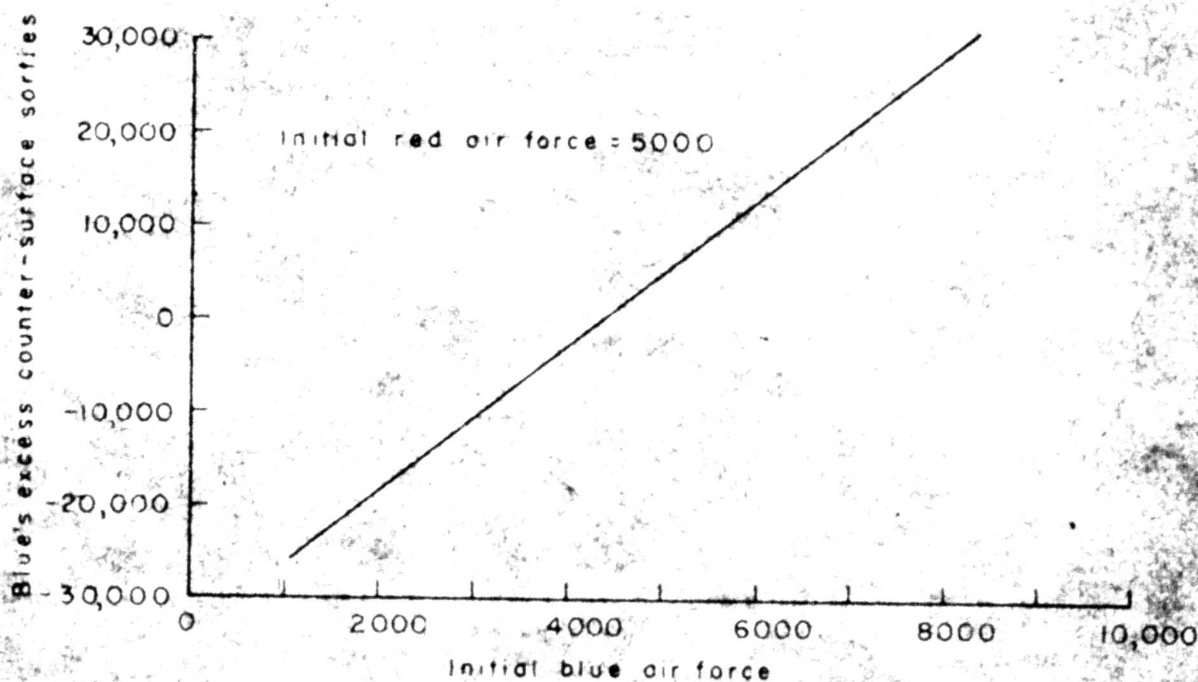


Fig. 7—Excess counter-surface sorties

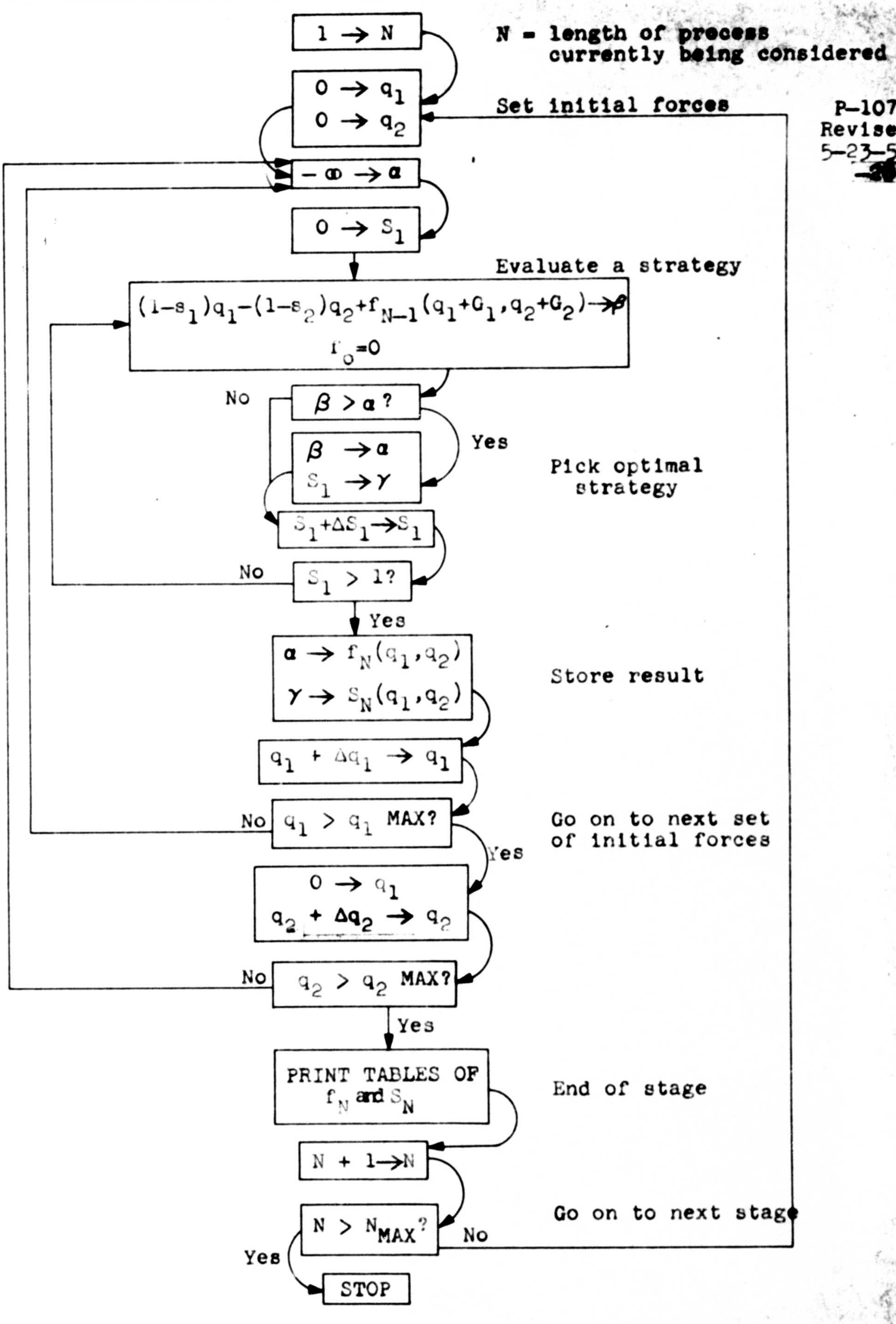


Fig. 8 — Flow-chart used for computational solution

BIBLIOGRAPHY

1. Bellman, Richard, Dynamic Programming, Princeton University Press, 1957.

2. ______________, The Theory of Dynamic Programming, Bull. Amer. Math. Soc., Vol 60, 1954, pp. 503-515.

3. Bellman, Richard, W. Fleming, and D. V. Widder, Variational Problems with Constraints, Annali di Matematica, Vol 16, 1956, pp. 301-323.

4. Johnson, Selmer, Best Exploration for Maximum is Fibonaccian, The RAND Corporation, Research Memorandum RM-1590, November 18, 1955.

5. Mengel, A. S., The RAND Corporation, unpublished paper.

BLANK PAGE